

PCT09

RAW SEQUENCE LISTING

DATE: 10/29/2001

PATENT APPLICATION: US/09/856,319

TIME: 11:30:53

Input Set : A:\UEMURA5.txt

Output Set: N:\CRF3\10292001\I856319.raw

3 <110> APPLICANT: UEMURA, Hidetoshi
4 OKUI, Akira
5 KOMINAMI, Katsuya
6 YAMAGUCHI, Nozomi
7 MITSUI, Shinichi
9 <120> TITLE OF INVENTION: NOVEL SERINE PROTEASE-BSSP5
11 <130> FILE REFERENCE: UEMURA=5
13 <140> CURRENT APPLICATION NUMBER: 09/856,319
14 <141> CURRENT FILING DATE: 2001-05-21
16 <150> PRIOR APPLICATION NUMBER: PCT/JP99/06473
17 <151> PRIOR FILING DATE: 1999-11-19
19 <150> PRIOR APPLICATION NUMBER: JP 347806/1998
20 <151> PRIOR FILING DATE: 1998-11-20
22 <160> NUMBER OF SEQ ID NOS: 33
24 <170> SOFTWARE: PatentIn version 3.1
26 <210> SEQ ID NO: 1
27 <211> LENGTH: 1149
28 <212> TYPE: DNA
29 <213> ORGANISM: Homo sapiens
31 <220> FEATURE:
32 <221> NAME/KEY: CDS
33 <222> LOCATION: (11)..(802)
34 <223> OTHER INFORMATION:
37 <400> SEQUENCE: 1
38 atctgccacg atg ttg ctg ctc agc ctg acc cta agc ctg gtt ctc ctc 49
39 Met Leu Leu Leu Ser Leu Thr Leu Ser Leu Val Leu Leu
40 1 5 10
42 ggc tcc tcc tgg ggc tgc ggc att cct gcc atc aaa ccg gca ctg agc 97
43 Gly Ser Ser Trp Gly Cys Gly Ile Pro Ala Ile Lys Pro Ala Leu Ser
44 15 20 25
46 ttc agc cag agg att gtc aac ggg gag aat gca gtg ttg ggc tcc tgg 145
47 Phe Ser Gln Arg Ile Val Asn Gly Glu Asn Ala Val Leu Gly Ser Trp
48 30 35 40 45
50 ccc tgg cag gtg tcc ctg cag gac agc agc ggc ttc cac ttc tgc ggt 193
51 Pro Trp Gln Val Ser Leu Gln Asp Ser Ser Gly Phe His Phe Cys Gly
52 50 55 60
54 ggt tct ctc atc agc cag tcc tgg gtg gtc act gct gcc cac tgc aat 241
55 Gly Ser Leu Ile Ser Gln Ser Trp Val Val Thr Ala Ala His Cys Asn
56 65 70 75
58 gtc agc cct ggc cgc cat ttt gtt gtc ctg ggc gag tat gac cga tca 289
59 Val Ser Pro Gly Arg His Phe Val Val Leu Gly Glu Tyr Asp Arg Ser
60 80 85 90
62 tca aac gca gag ccc ttg cag gtt ctg tcc gtc tct cgg gcc att aca 337
63 Ser Asn Ala Glu Pro Leu Gln Val Leu Ser Val Ser Arg Ala Ile Thr
64 95 100 105
66 cac cct agc tgg aac tct acc acc atg aac aat gac gtg acg ctg ctg 385
67 His Pro Ser Trp Asn Ser Thr Thr Met Asn Asn Asp Val Thr Leu Leu

ENTERED

p. 5

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/856,319

DATE: 10/29/2001

TIME: 11:30:53

Input Set : A:\UEMURA5.txt

Output Set: N:\CRF3\10292001\I856319.raw

```

68 110          115          120          125
70 aag ctc gcc tcg cca gcc cag tac aca aca cgc atc tcg cca gtt tgc      433
71 Lys Leu Ala Ser Pro Ala Gln Tyr Thr Thr Arg Ile Ser Pro Val Cys
72          130          135          140
74 ctg gca tcc tca aac gag gct ctg act gaa ggc ctc acg tgt gtc acc      481
75 Leu Ala Ser Ser Asn Glu Ala Leu Thr Glu Gly Leu Thr Cys Val Thr
76          145          150          155
78 acc ggc tgg ggt cgc ctc agt ggc gtg ggc aat gtg aca cca gca cat      529
79 Thr Gly Trp Gly Arg Leu Ser Gly Val Gly Asn Val Thr Pro Ala His
80          160          165          170
82 ctg cag cag gtg gct ttg ccc ctg gtc act gtg aat cag tgc cgg cag      577
83 Leu Gln Gln Val Ala Leu Pro Leu Val Thr Val Asn Gln Cys Arg Gln
84          175          180          185
86 tac tgg gac tca agt atc act gac tcc atg atc tgt gca ggt ggc gca      625
87 Tyr Trp Asp Ser Ser Ile Thr Asp Ser Met Ile Cys Ala Gly Gly Ala
88 190          195          200          205
90 ggt gcc tcc tcg tgc cag ggt gac tcc gga ggc cct ctt gtc tgc cag      673
91 Gly Ala Ser Ser Cys Gln Gly Asp Ser Gly Gly Pro Leu Val Cys Gln
92          210          215          220
94 aag gga aac aca tgg gtg ctt att ggt att gtc tcc tgg ggc acc aaa      721
95 Lys Gly Asn Thr Trp Val Leu Ile Gly Ile Val Ser Trp Gly Thr Lys
96          225          230          235
98 aac tgc aat gtg cgc gca cct gct gtg tat act cga gtt agc aag ttc      769
99 Asn Cys Asn Val Arg Ala Pro Ala Val Tyr Thr Arg Val Ser Lys Phe
100          240          245          250
102 agc acc tgg atc aac cag gtc ata gcc tac aac tgagctcacc acaggccctc      822
103 Ser Thr Trp Ile Asn Gln Val Ile Ala Tyr Asn
104          255          260
106 cccagctcaa cccatttaaa ggacccaggc cctgtcccat catgcattca tgtctgtctt      882
108 cctggctcag gagaaagaag aggctgttga gggctccgact ccctacttgg acttctggca      942
110 cagaaggggc tgagtgaactc cttgagtagc agtggctctt cctagagtag ccatgccgtg      1002
112 gccggggccc ccacccctcc tccagggcaa ccccttggtc ctacagcaag aagccagaac      1062
114 tgttggaatg aatggcagcc ctccctggag aggcagcctg tttactgaat acagaggata      1122
116 cgtttacaaa aaaaaaaaaa aaaaaaaa      1149
119 <210> SEQ ID NO: 2
120 <211> LENGTH: 264
121 <212> TYPE: PRT
122 <213> ORGANISM: Homo sapiens
124 <400> SEQUENCE: 2
126 Met Leu Leu Leu Ser Leu Thr Leu Ser Leu Val Leu Leu Gly Ser Ser
127 1          5          10          15
130 Trp Gly Cys Gly Ile Pro Ala Ile Lys Pro Ala Leu Ser Phe Ser Gln
131          20          25          30
134 Arg Ile Val Asn Gly Glu Asn Ala Val Leu Gly Ser Trp Pro Trp Gln
135          35          40          45
138 Val Ser Leu Gln Asp Ser Ser Gly Phe His Phe Cys Gly Gly Ser Leu
139          50          55          60
142 Ile Ser Gln Ser Trp Val Val Thr Ala Ala His Cys Asn Val Ser Pro
143 65          70          75          80

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/856,319

DATE: 10/29/2001

TIME: 11:30:53

Input Set : A:\UEMURA5.txt

Output Set: N:\CRF3\10292001\I856319.raw

```

146 Gly Arg His Phe Val Val Leu Gly Glu Tyr Asp Arg Ser Ser Asn Ala
147      85      90      95
150 Glu Pro Leu Gln Val Leu Ser Val Ser Arg Ala Ile Thr His Pro Ser
151      100      105      110
154 Trp Asn Ser Thr Thr Met Asn Asn Asp Val Thr Leu Leu Lys Leu Ala
155      115      120      125
158 Ser Pro Ala Gln Tyr Thr Thr Arg Ile Ser Pro Val Cys Leu Ala Ser
159      130      135      140
162 Ser Asn Glu Ala Leu Thr Glu Gly Leu Thr Cys Val Thr Thr Gly Trp
163 145      150      155      160
166 Gly Arg Leu Ser Gly Val Gly Asn Val Thr Pro Ala His Leu Gln Gln
167      165      170      175
170 Val Ala Leu Pro Leu Val Thr Val Asn Gln Cys Arg Gln Tyr Trp Asp
171      180      185      190
174 Ser Ser Ile Thr Asp Ser Met Ile Cys Ala Gly Gly Ala Gly Ala Ser
175      195      200      205
178 Ser Cys Gln Gly Asp Ser Gly Gly Pro Leu Val Cys Gln Lys Gly Asn
179      210      215      220
182 Thr Trp Val Leu Ile Gly Ile Val Ser Trp Gly Thr Lys Asn Cys Asn
183 225      230      235      240
186 Val Arg Ala Pro Ala Val Tyr Thr Arg Val Ser Lys Phe Ser Thr Trp
187      245      250      255
190 Ile Asn Gln Val Ile Ala Tyr Asn
191      260
194 <210> SEQ ID NO: 3
195 <211> LENGTH: 834
196 <212> TYPE: DNA
197 <213> ORGANISM: mus sp.
199 <220> FEATURE:
200 <221> NAME/KEY: CDS
201 <222> LOCATION: (33)..(824)
202 <223> OTHER INFORMATION:
205 <400> SEQUENCE: 3
206 gaccatctca acaccattcc ttatttgtca ca atg cta ctg ctc agc cta acc      53
207      Met Leu Leu Leu Ser Leu Thr
208      1      5
210 ctt agc ctg gtc ctc ctt ggc tcc tcc tgg ggc tgt ggt gtt cct gcc      101
211 Leu Ser Leu Val Leu Leu Gly Ser Ser Trp Gly Cys Gly Val Pro Ala
212      10      15      20
214 atc acg cct gca ctg agc tac aat cag aga att gtc aac ggg gag aat      149
215 Ile Thr Pro Ala Leu Ser Tyr Asn Gln Arg Ile Val Asn Gly Glu Asn
216      25      30      35
218 gca gtg cca ggc tcc tgg ccc tgg cag gtg tct ctc cag gat aac acc      197
219 Ala Val Pro Gly Ser Trp Pro Trp Gln Val Ser Leu Gln Asp Asn Thr
220 40      45      50      55
222 ggc ttc cac ttc tgc ggt ggt tct ctc atc agt ccg aac tgg gtg gtc      245
223 Gly Phe His Phe Cys Gly Gly Ser Leu Ile Ser Pro Asn Trp Val Val
224      60      65      70
226 acg gct gcc cac tgc caa gtc acg cct gga cgc cac ttt gtc gtt ttg      293

```

RAW SEQUENCE LISTING

DATE: 10/29/2001

PATENT APPLICATION: US/09/856,319

TIME: 11:30:53

Input Set : A:\UEMURA5.txt

Output Set: N:\CRF3\10292001\I856319.raw

```

227 Thr Ala Ala His Cys Gln Val Thr Pro Gly Arg His Phe Val Val Leu
228      75      80      85
230 gga gaa tat gac cga tct tcc aat gct gaa cct gtg cag gtc ctc tcg      341
231 Gly Glu Tyr Asp Arg Ser Ser Asn Ala Glu Pro Val Gln Val Leu Ser
232      90      95      100
234 atc gca agg gcc atc aca cac cct aac tgg aac gcc aac acc atg aac      389
235 Ile Ala Arg Ala Ile Thr His Pro Asn Trp Asn Ala Asn Thr Met Asn
236      105      110      115
238 aat gac ctg act ctc ctg aag ctt gcc tcg cca gcc cgg tac aca gca      437
239 Asn Asp Leu Thr Leu Leu Lys Leu Ala Ser Pro Ala Arg Tyr Thr Ala
240 120      125      130      135
242 caa gtc tca cca gtc tgc ctg gct tcc aca aac gag gca ctg cct tcg      485
243 Gln Val Ser Pro Val Cys Leu Ala Ser Thr Asn Glu Ala Leu Pro Ser
244      140      145      150
246 ggg ctc acc tgt gtc acc act ggc tgg ggc cga atc agt ggt gtg ggc      533
247 Gly Leu Thr Cys Val Thr Thr Gly Trp Gly Arg Ile Ser Gly Val Gly
248      155      160      165
250 aat gtg aca cca gct cgc ctg cag caa gtt gtt cta ccc ctg gtc act      581
251 Asn Val Thr Pro Ala Arg Leu Gln Gln Val Val Leu Pro Leu Val Thr
252      170      175      180
254 gtg aat cag tgt cgg cag tac tgg ggt gca cgc att acc gat gcc atg      629
255 Val Asn Gln Cys Arg Gln Tyr Trp Gly Ala Arg Ile Thr Asp Ala Met
256      185      190      195
258 ata tgt gca ggt ggc tca ggc gcc tcc tca tgt cag ggt gac tca gga      677
259 Ile Cys Ala Gly Gly Ser Gly Ala Ser Ser Cys Gln Gly Asp Ser Gly
260 200      205      210      215
262 ggc cct ctt gtc tgc cag aag gga aac acc tgg gtg ctt att ggg att      725
263 Gly Pro Leu Val Cys Gln Lys Gly Asn Thr Trp Val Leu Ile Gly Ile
264      220      225      230
266 gtc tcc tgg ggc act aag aac tgc aac ata caa gca ccg gcc atg tac      773
267 Val Ser Trp Gly Thr Lys Asn Cys Asn Ile Gln Ala Pro Ala Met Tyr
268      235      240      245
270 act cgg gtc agc aag ttc agt acc tgg atc aac caa gtc atg gcc tac      821
271 Thr Arg Val Ser Lys Phe Ser Thr Trp Ile Asn Gln Val Met Ala Tyr
272      250      255      260
274 aac taaactgtcc      834
275 Asn
279 <210> SEQ ID NO: 4
280 <211> LENGTH: 264
281 <212> TYPE: PRT
282 <213> ORGANISM: mus sp.
284 <400> SEQUENCE: 4
286 Met Leu Leu Leu Ser Leu Thr Leu Ser Leu Val Leu Leu Gly Ser Ser
287 1      5      10      15
290 Trp Gly Cys Gly Val Pro Ala Ile Thr Pro Ala Leu Ser Tyr Asn Gln
291      20      25      30
294 Arg Ile Val Asn Gly Glu Asn Ala Val Pro Gly Ser Trp Pro Trp Gln
295      35      40      45
298 Val Ser Leu Gln Asp Asn Thr Gly Phe His Phe Cys Gly Gly Ser Leu

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/856,319

DATE: 10/29/2001

TIME: 11:30:53

Input Set : A:\UEMURA5.txt

Output Set: N:\CRF3\10292001\I856319.raw

```

299      50      55      60
302 Ile Ser Pro Asn Trp Val Val Thr Ala Ala His Cys Gln Val Thr Pro
303 65      70      75      80
306 Gly Arg His Phe Val Val Leu Gly Glu Tyr Asp Arg Ser Ser Asn Ala
307      85      90      95
310 Glu Pro Val Gln Val Leu Ser Ile Ala Arg Ala Ile Thr His Pro Asn
311      100     105     110
314 Trp Asn Ala Asn Thr Met Asn Asn Asp Leu Thr Leu Leu Lys Leu Ala
315      115     120     125
318 Ser Pro Ala Arg Tyr Thr Ala Gln Val Ser Pro Val Cys Leu Ala Ser
319      130     135     140
322 Thr Asn Glu Ala Leu Pro Ser Gly Leu Thr Cys Val Thr Thr Gly Trp
323 145     150     155     160
326 Gly Arg Ile Ser Gly Val Gly Asn Val Thr Pro Ala Arg Leu Gln Gln
327      165     170     175
330 Val Val Leu Pro Leu Val Thr Val Asn Gln Cys Arg Gln Tyr Trp Gly
331      180     185     190
334 Ala Arg Ile Thr Asp Ala Met Ile Cys Ala Gly Gly Ser Gly Ala Ser
335      195     200     205
338 Ser Cys Gln Gly Asp Ser Gly Gly Pro Leu Val Cys Gln Lys Gly Asn
339      210     215     220
342 Thr Trp Val Leu Ile Gly Ile Val Ser Trp Gly Thr Lys Asn Cys Asn
343 225     230     235     240
346 Ile Gln Ala Pro Ala Met Tyr Thr Arg Val Ser Lys Phe Ser Thr Trp
347      245     250     255
350 Ile Asn Gln Val Met Ala Tyr Asn

```

351 260

354 <210> SEQ ID NO: 5

355 <211> LENGTH: 99

356 <212> TYPE: DNA

357 <213> ORGANISM: Artificial Sequence

359 <220> FEATURE:

360 <223> OTHER INFORMATION: Designed oligonucleotide to construct plasmid pSecTrypHis

362 <400> SEQUENCE: 5

363 aagcttggct agcaacacca tgaatctact cctgatocctt acctttgttg ctgctgctgt 60

365 tgctgcccc tttgacgacg atgacaagga tccgaattc 99

368 <210> SEQ ID NO: 6

369 <211> LENGTH: 99

370 <212> TYPE: DNA

371 <213> ORGANISM: Artificial Sequence

373 <220> FEATURE:

374 <223> OTHER INFORMATION: Designed oligonucleotide to construct plasmid pSecTrypHis

376 <400> SEQUENCE: 6

377 gaattcggat ccttgatcat gtcgtcaaag ggggcagcaa cagcagcagc aacaaaggta 60

379 aggatcagga gtagattcat ggtgttgcta gccaaagott 99

382 <210> SEQ ID NO: 7

383 <211> LENGTH: 15

384 <212> TYPE: DNA

385 <213> ORGANISM: Artificial Sequence

Use of n and / or Xaa has been detected in the
Sequence Listing. Review the Sequence Listing
to ensure a corresponding explanation is present
in the <220> to <223> fields of each sequence
using n or Xaa.

Use of n and / or Xaa has been detected in the
Sequence Listing. Review the Sequence Listing
to ensure a corresponding explanation is present
in the <220> to <223> fields of each sequence
using n or Xaa

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/856,319

DATE: 10/29/2001

TIME: 11:30:54

Input Set : A:\UEMURA5.txt

Output Set: N:\CRF3\10292001\I856319.raw

L:495 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:14

L:520 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:15